USS Lead Soil Contamination Site

East Chicago, IN

FIELDS Statistical Evaluation Summary

12 February 2007

Data Sources

Data included in these analyses were compiled over five (5) separate sampling events in which surface soil samples were collected and analyzed for lead (Pb) concentrations on-site or near the USS Lead Site. Table 1 summarizes the results of each sampling event. There are a couple of caveats associated with compiling the data from different sampling events: 1) Most sampling events included surface soil samples at the 0-6 inch soil depth interval. The only exception to this method occurs in the EPA 2006 sampling event in which samples were collected at the 0-1 inch and 1-6 inch soil depth interval. Therefore, the compiled dataset is composed of an inconsistent sampling method as only the 0-1 inch depth interval was used for the EPA 2006 sampling event. 2) Most sampling events report laboratory analytic results only (i.e., no XRF values). The only exception to this occurs in the RCRA 2003 sampling event, where XRF was used to analyze all surface soil samples. A subset of these samples (10%) was submitted for laboratory analytical analyses. For these 2003 RCRA samples, regression was used to determine the linear relationship between XRF values and laboratory analytic results. The resulting regression equation was used to adjust the XRF values for those 2003 RCRA samples that were not submitted for laboratory analysis.

A total of 258 surface soil samples were analyzed for the entire compiled dataset. Because the ENTACT 1999 samples were primarily collected on-site near the perimeter of the USS Lead site boundary, these samples are referred to as "On-Site Samples" ($N = 88$). The remaining samples are referred to as "Off-Site Samples" ($N = 170$). All analyses and figures reported herein are created with and without the influence of the on-site samples.

Correlation & Regression

Correlation and regression analyses were conducted to determine the relationship between surface soil lead (Pb) concentrations and distance from the USS Lead site boundary. Pearson (parametric) and Spearman (non-parametric) correlation coefficients are reported in all analyses. All analyses are performed using un-transformed lead concentrations and log_{10}-transformed lead concentrations. Results of these analyses using the entire compiled dataset ($N = 258$) are shown in Table 2. Results of these analyses using only the off-site samples ($N = 170$) are shown in Table 3. All correlation and regression analyses indicate that a negative association exists between the lead concentrations and distance from the site (Figures 1-4). Overall, analyses using log_{10}-transformed concentrations revealed stronger relationships than analyses using un-transformed concentrations. The strongest relationship existed for log_{10}-transformed off-site samples (Table 3B; Figure 4).

Categorical Distance Measurements

All samples were assigned to one of five categories based upon distance from the USS Lead site boundary. Distance categories are as follows: 0-250m, 250-500m, 500-750m, 750-1000m, & >1000m. Table 4 summarizes the lead concentrations within the distance categories for the entire compiled dataset ($N = 258$) and Table 5 summarizes only the off-site samples within the distance categories ($N = 170$). These tables indicate that the average lead concentration decreases with increasing distance from the USS Lead site boundary. Furthermore, Figures 5 and 6 also show the negative relationship between average lead concentrations and distance for the entire compiled dataset and off-site samples, respectively. A generalized linear model (GENMOD) indicated that distance from the USS Lead site boundary significantly influences the average surface soil lead concentrations for the entire compiled dataset and off-site samples.

Table 1. Sources of information used in analyses to associate surface soil lead (Pb) concentrations with distance from the USS Lead Soil Contamination Site, East Chicago, IN. Results have been obtained using various methods and analytical procedures.

Sampling Event	N	Soil Interval	Analytical Method	Average Results (ppm)	Standard Deviation (ppm)	Median (ppm)	Min (ppm)	Max (ppm)
ENTACT, 1999	88	0 - 6 Inches	LAB	14049.5	21216.1	3750.0	150.0	110000.0
EPA / IDEM, 2002	23	0 - 6 Inches	LAB	545.6	1379.9	210.0	17.0	6700.0
EPA, 1985	36	0 - 6 Inches	LAB	1069.5	2317.0	280.0	30.0	11000.0
EPA, 2006	28	0 - 1 Inches	LAB	866.5	458.0	860.0	92.0	2300.0
RCRA, 2003	83	0 - 6 Inches	XRF (adjusted)[1]	940.6	465.6	787.0	538.1	4030.0
TOTAL	258	0 - 6 Inches	LAB	5386.6	13870.0	900.6	17.0	110000.0

[1] XRF values were adjusted using a linear relationship with laboratory analytical results

Table 2. Correlation and regression analyses for on-site and off-site samples at the USS Lead Soil Contamination Site, East Chicago, IN. These analyses determined if a relationship existed between surface soil lead (Pb) concentrations and distance from the USS Lead Site. A) Raw (untransformed) data. B) Log10-transformed lead (Pb) concentrations. Results are obtained from different sources using a variety of methods and analytical procedures. Total number of on-site and off-site samples analyzed $N = 258$.

A) Raw (Untransformed) Results*

CORRELATIONS	Coefficient	P-Value
Pearson (r)	-0.301	< 0.001
Spearman (r)	-0.670	< 0.001

LINEAR REGRESSION				
Treatment	Df	F-Value	P-Value	Model R^2
Model	1, 257	25.6	< 0.001	0.087

Variable	Df	Coefficient	t-Value	P-Value
Intercept	1, 257	9030.7	8.2	< 0.001
Distance	1, 257	-7.0	-5.1	

B) Log10-Transformed Lead Concentrations**

CORRELATIONS	Coefficient	P-Value
Pearson (r)	-0.643	< 0.001
Spearman (r)	-0.670	< 0.001

LINEAR REGRESSION				
Treatment	Df	F-Value	P-Value	Model R^2
Model	1, 257	180.8	< 0.001	0.412

Variable	Df	Coefficient	t-Value	P-Value
Intercept	1, 257	3.4	77.1	< 0.001
Distance	1, 257	-0.0008	-13.5	< 0.001

*   See Figure 1
** See Figure 2
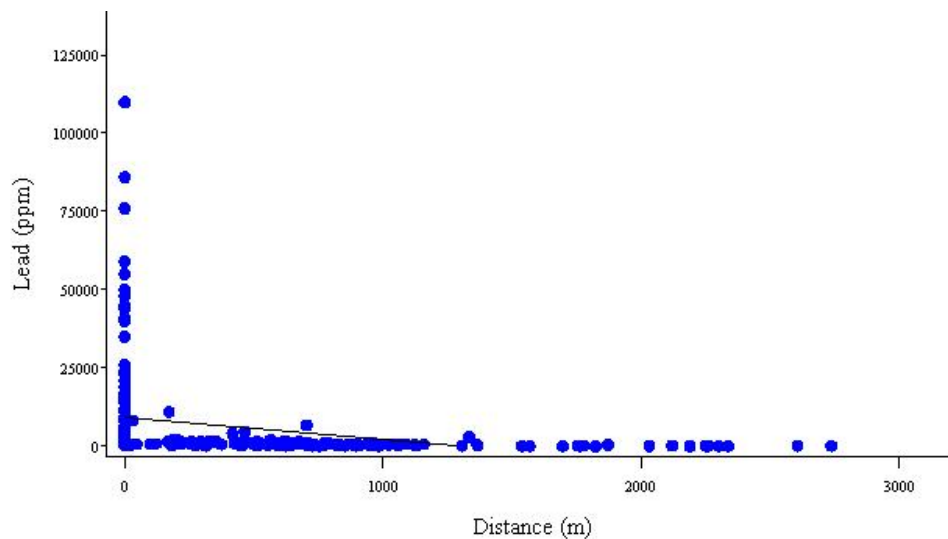
Figure 1. Scatter plot showing the relationship between distance from the USS Lead Site and lead (Pb) concentration for on-site and off-site samples ($N = 258$). Results (dots) are shown in relation to predicted values (black regression line).
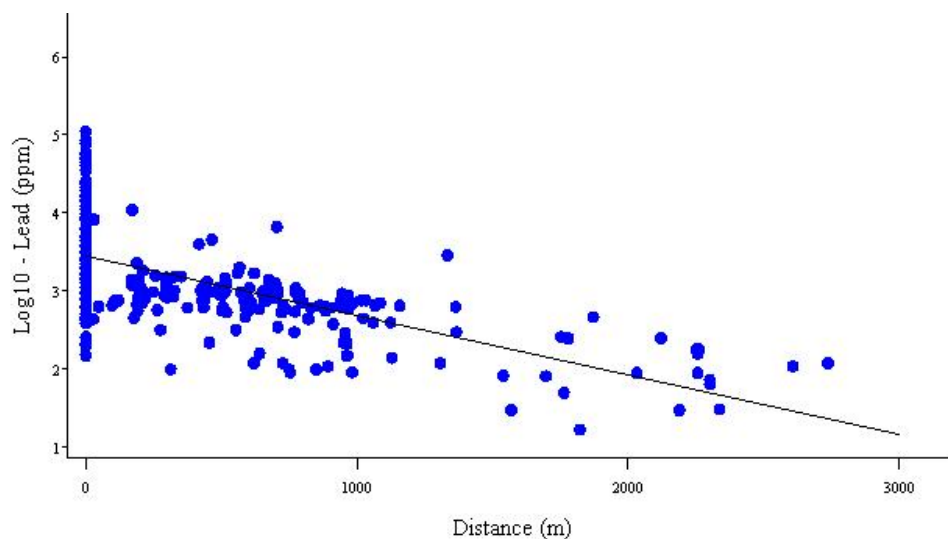
Figure 2. Relationship between sample distance from the USS Lead Site and log10-transformed lead (Pb) concentration for on-site and off-site samples ($N = 258$). Log10-transformed values (dots) are shown in relation to predicted values (black regression line).

Table 3. Correlation and regression analyses for off-site samples at the USS Lead Soil Contamination Site, East Chicago, IN. These analyses determined if a relationship existed between surface soil lead (Pb) concentrations and distance from the USS Lead Site. A) Raw (untransformed) data. B) Log_{10}-transformed lead (Pb) concentrations. ENTACT (1999) samples were presumably collected on-site and are omitted from this table. Results are obtained from different sources using a variety of methods and analytical procedures. Total number of off-site samples analyzed $N = 170$.

A) Raw (Untransformed) Results[+]

CORRELATIONS			LINEAR REGRESSION				
	Coefficient	*P*-Value	Treatment	Df	*F*-Value	*P*-Value	Model R^2
Pearson (r)	-0.328	< 0.001	Model	1, 169	20.22	< 0.001	0.102
Spearman (r)	-0.589	< 0.001					
			Variable	Df	Coefficient	*t*-Value	*P*-Value
			Intercept	1, 169	1460.1	9.54	< 0.001
			Distance	1, 169	-0.709	-4.50	< 0.001

B) Log10-Transformed Lead Concentrations[++]

| CORRELATIONS			LINEAR REGRESSION				
	Coefficient	*P*-Value	Treatment	Df	*F*-Value	*P*-Value	Model R^2
Pearson (r)	-0.659	< 0.001	Model	1, 169	128.93	< 0.001	0.431
Spearman (r)	-0.589	< 0.001					
			Variable	Df	Coefficient	*t*-Value	*P*-Value
			Intercept	1, 169	0.045	70.45	< 0.001
			Distance	1, 169	-0.00052	-11.35	< 0.001

[+] See Figure 3
[++] See Figure 4
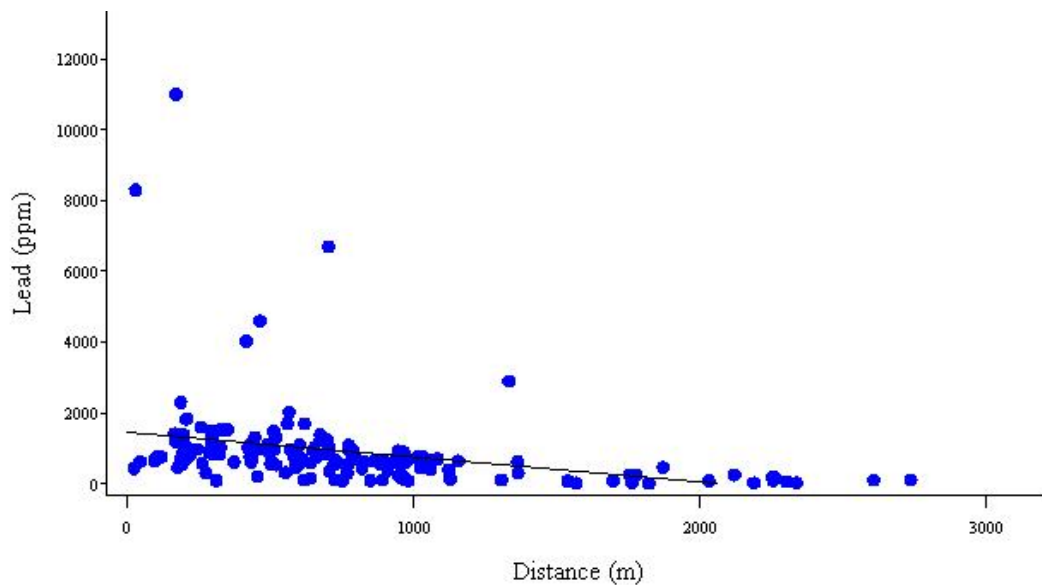
Figure 3. Scatter plot showing the relationship between distance from the USS Lead Site and lead (Pb) concentration for off-site samples ($N = 170$). ENTACT (1999) samples were presumably collected on-site and are omitted from this analysis. Results (dots) are shown in relation to predicted values (black regression line).
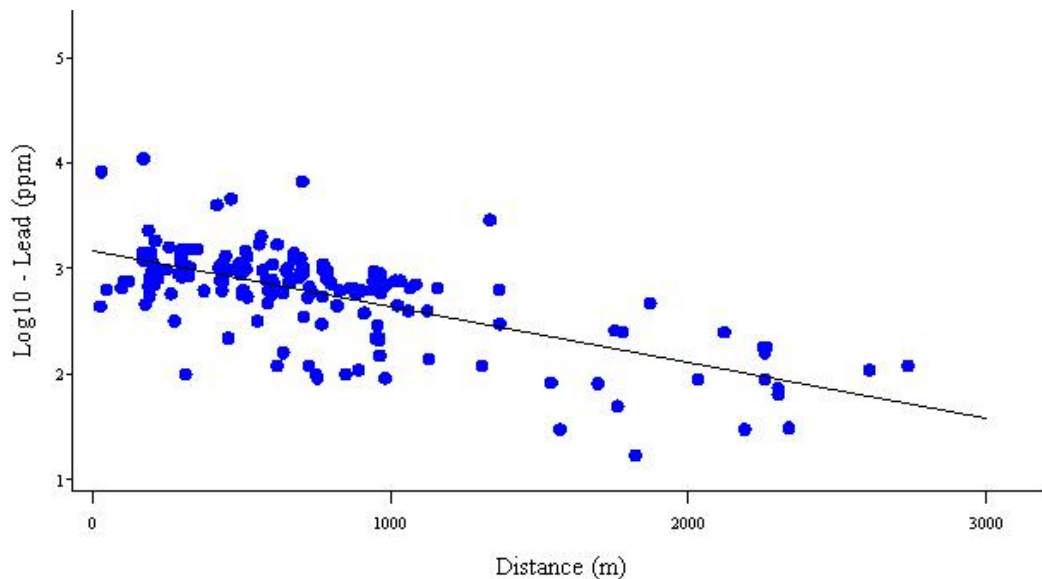
Figure 4. Relationship between sample distance from the USS Lead Site and log10-transformed lead (Pb) concentration for off-site samples ($N = 170$). ENTACT (1999) samples were presumably collected on-site and are omitted from this analysis. Log_{10}-transformed values (dots) are shown in relation to predicted values (black regression line).

Table 4. Summary of surface soil lead (Pb) concentrations analyzed from 1985 - 2006 along distance categories. On-site and off-site samples are included. Note: Results are obtained from different sources using various analytical methods. Total number of on-site and off-site samples $N = 258$.

Distance Category	N	Average Concentration (ppm)	Standard Deviation (ppm)	Standard Error (ppm)	Median	Min	Max	Upper 95% Confidence Limit
0 - 250 m	115	11139.2	19304.2	1800.1	2000.0	150.0	110000.0	14124.4
250 - 500 m	30	1178.4	931.3	170.0	1008.4	100.0	4600.0	1467.3
500 - 750 m	38	1006.0	1045.1	169.5	802.9	120.0	6700.0	1292.1
750 - 1000 m	42	562.9	291.3	44.9	638.1	92.0	1100.0	638.5
>1000 m	33	349.1	517.0	90.0	180.0	17.0	2900.0	501.5
TOTAL	258	5386.6	13870.0	863.5	900.6	17.0	110000.0	6812.1

GENMOD (Normal Regression):

$X^2 = 38.7$; $P < 0.001$
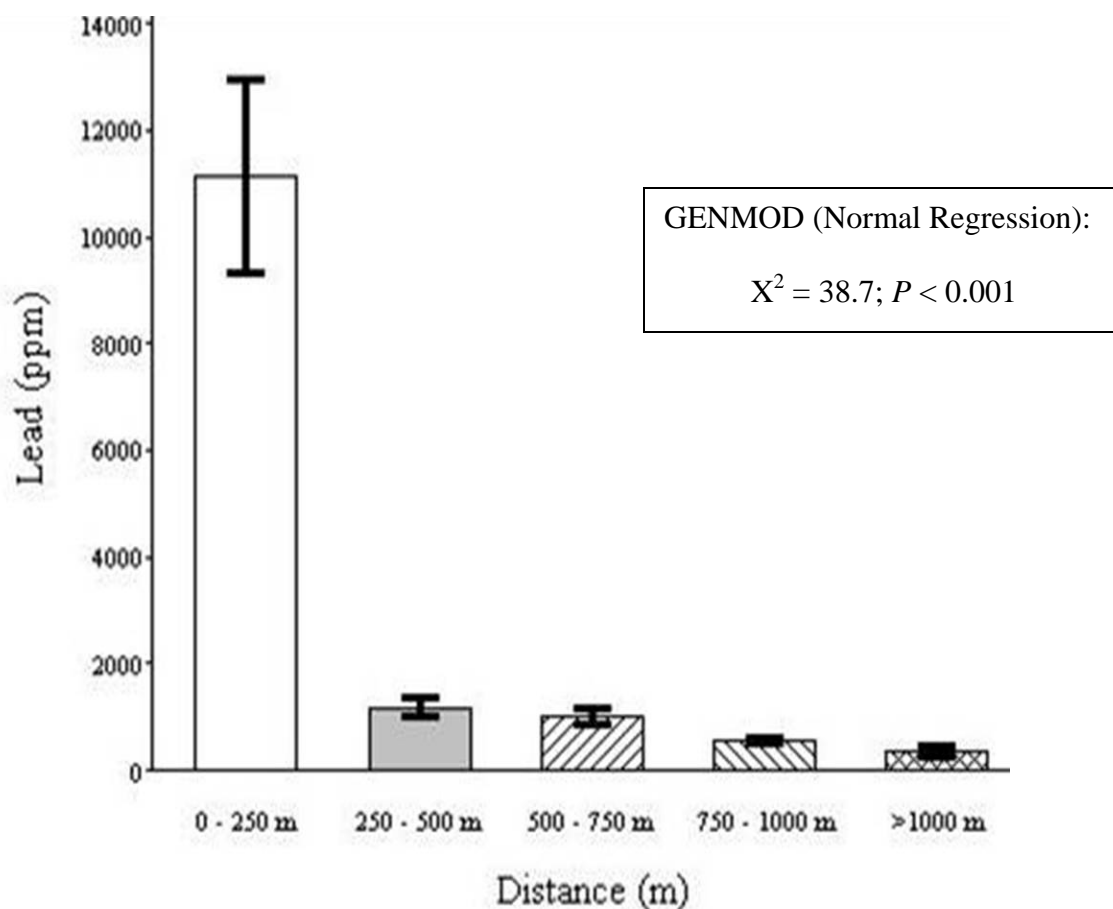
Figure 5. Average lead (Pb) concentration ($\pm$ 1 SE) of on-site and off-site soil samples with increasing distance from the USS Lead Site. Total number of on-site and off-site samples $N = 258$.

Table 5. Summary of off-site surface soil lead (Pb) concentrations analyzed from 1985 - 2006 along distance categories. ENTACT (1999) samples were presumably collected on-site and are omitted from this table. Note: Results are obtained from different sources using various analytical methods. Total number of off-site samples in this analysis $N = 170$.

Distance Category	N	Average Concentration (ppm)	Standard Deviation (ppm)	Standard Error (ppm)	Median	Min	Max	Upper 95% Confidence Limit
0 - 250 m	27	1653.5	2376.6	457.4	982.9	440.0	11000.0	2433.6
250 - 500 m	30	1178.4	931.3	170.0	1008.4	100.0	4600.0	1467.3
500 - 750 m	38	1006.0	1045.0	169.5	802.8	120.0	6700.0	1292.0
750 - 1000 m	42	562.9	291.3	44.9	638.1	92.0	1100.0	638.5
>1000 m	33	349.1	517.0	90.0	180.0	17.0	2900.0	501.5
TOTAL	170	902.3	1233.9	94.6	711.0	17.0	11000.0	1058.8

GENMOD (Normal Regression):

$X^2 = 23.2; P < 0.001$
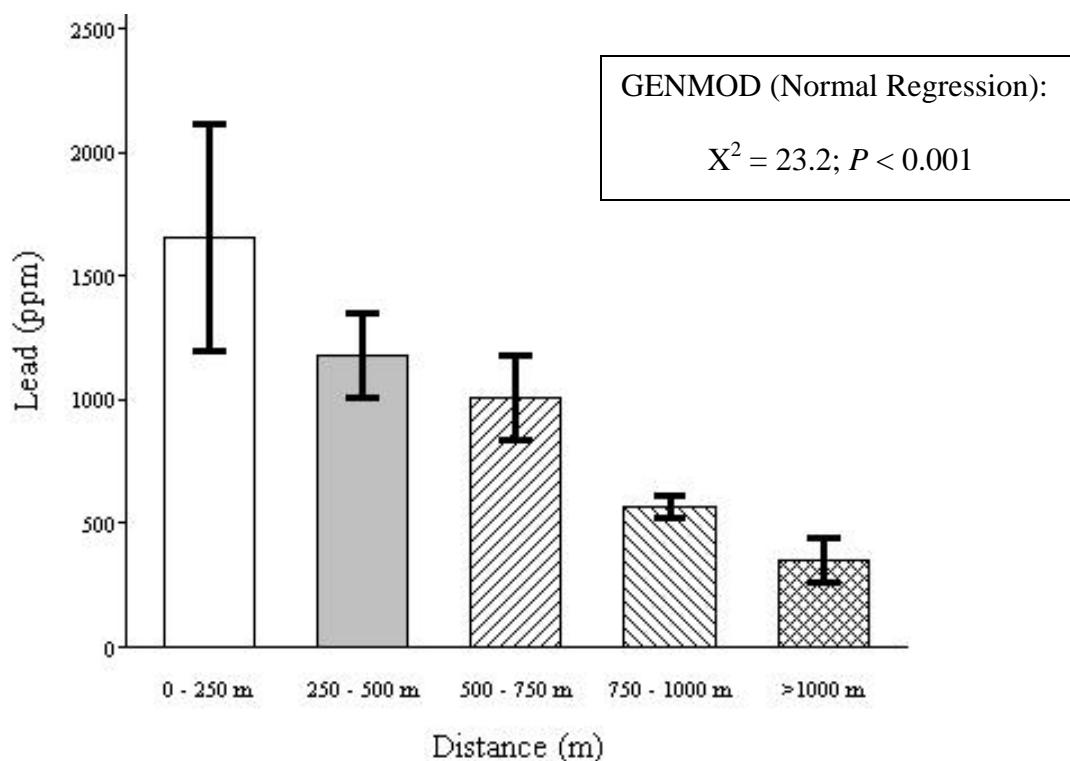
Figure 6. Average lead (Pb) concentration ($\pm$ 1 SE) of off-site soil samples with increasing distance from the USS Lead Site. ENTACT (1999) samples were presumably collected on-site and are omitted from this figure. Total number of off-site samples $N = 170$.

Addendum 1

Q-Q Plots

Q-Q Plots by Distance

Normal Q-Q Plots

All Distances / Events
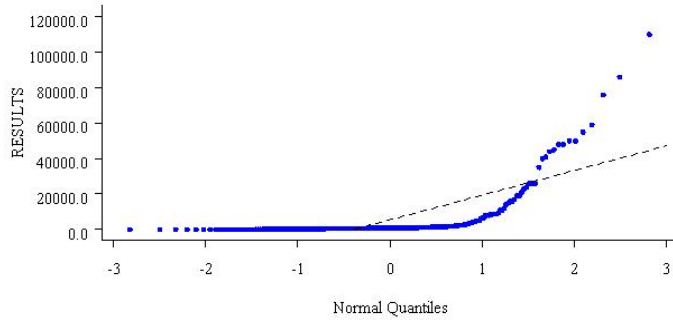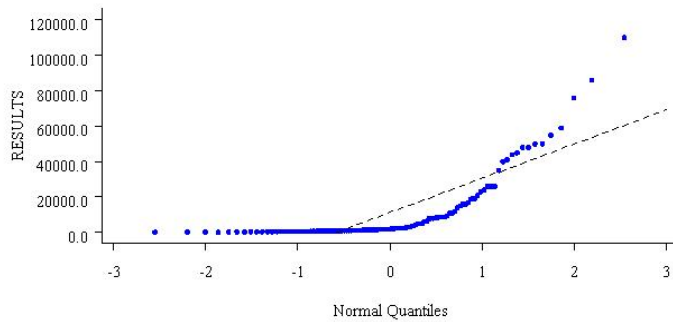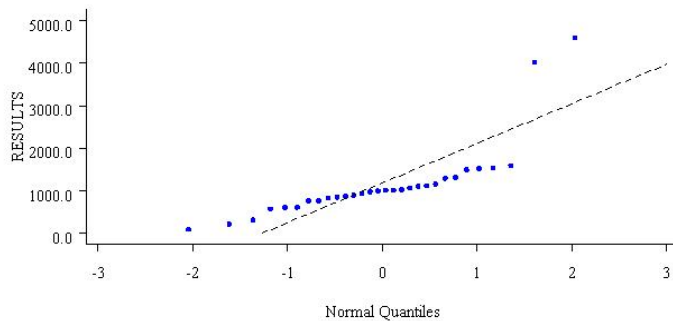
0 - 250 m

250 - 500 m

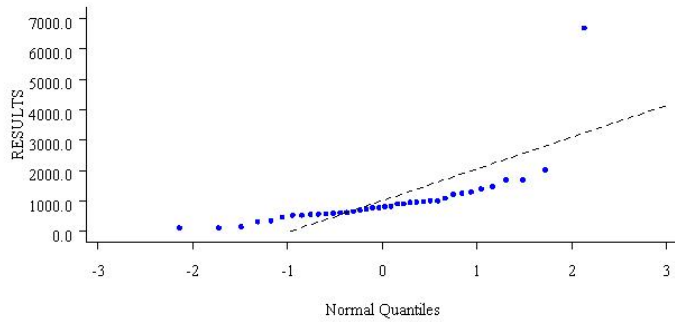## Log-Normal Q-Q Plots

All Distances / Events
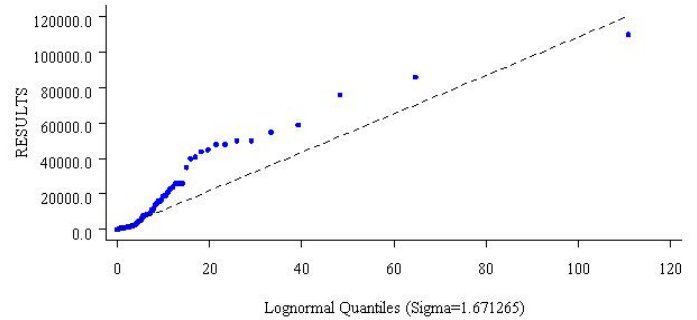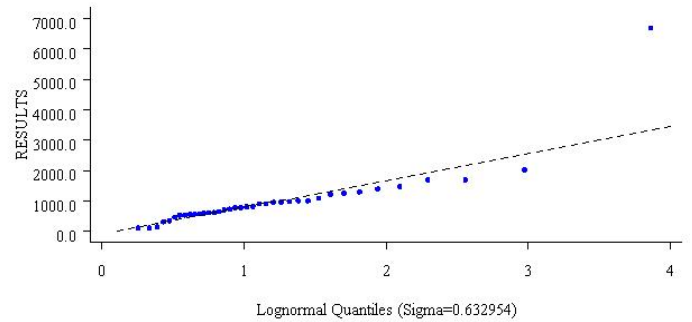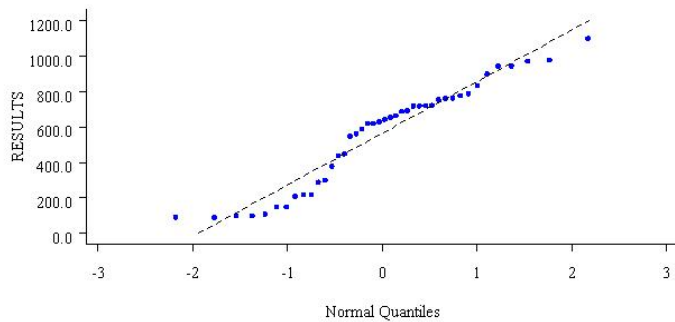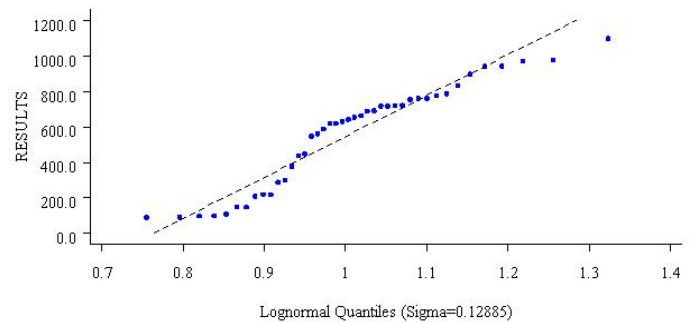

0 - 250 m
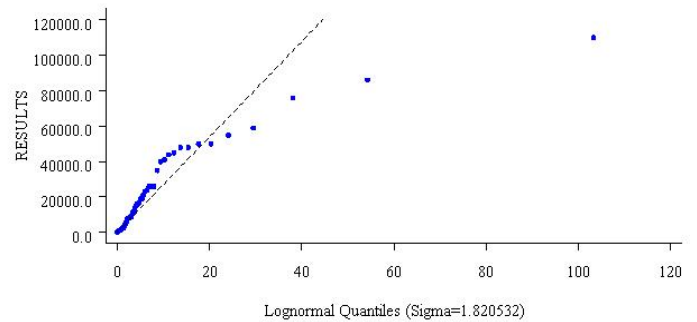
250 - 500 m
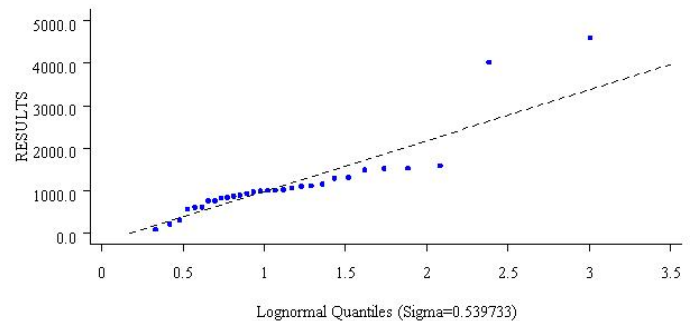
Normal Q-Q Plots

Log-Normal Q-Q Plots

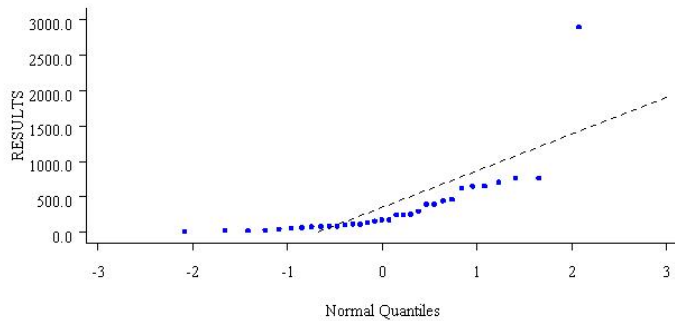### 500 - 750 m
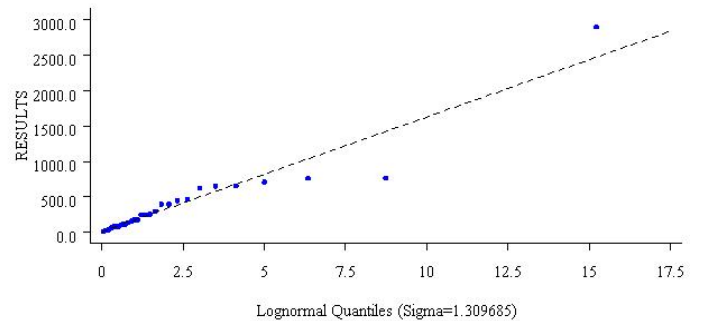

500 - 750 m

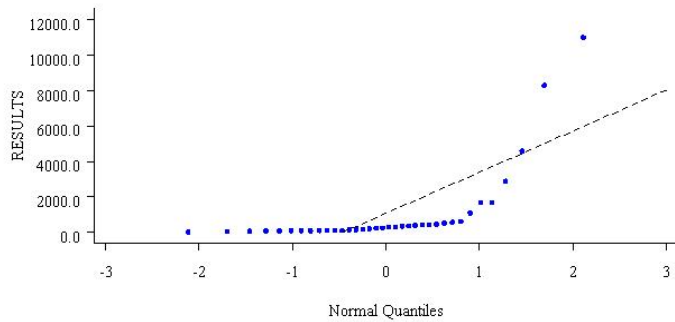
750 - 1000 m

750 - 1000 m

> 1000 m

> 1000 m

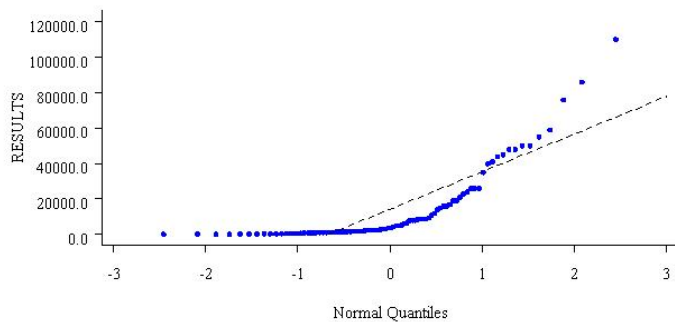# Q-Q Plots by Sample Event
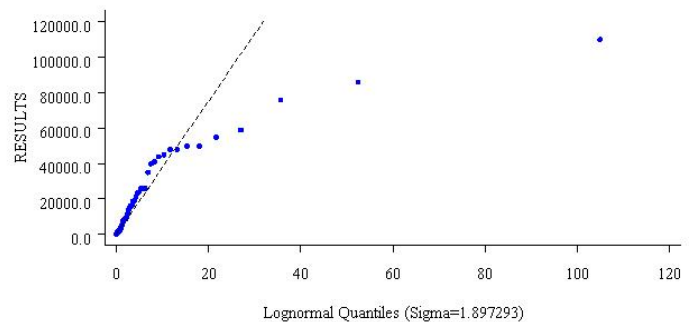
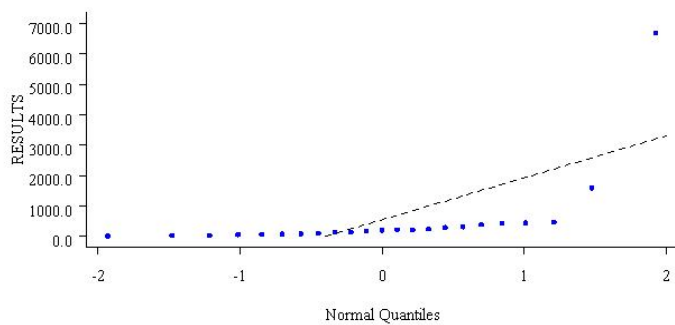## Normal Q-Q Plots

Log-Normal Q-Q Plots

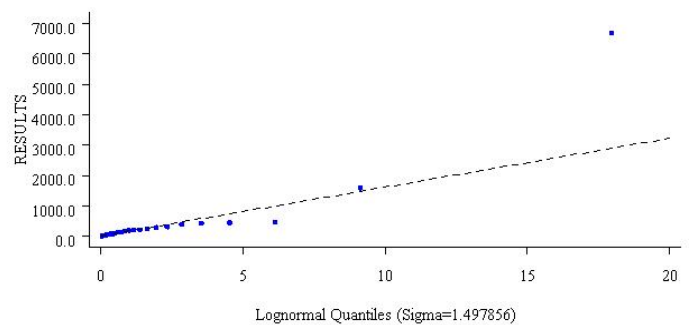### EPA, 1985

EPA, 1985
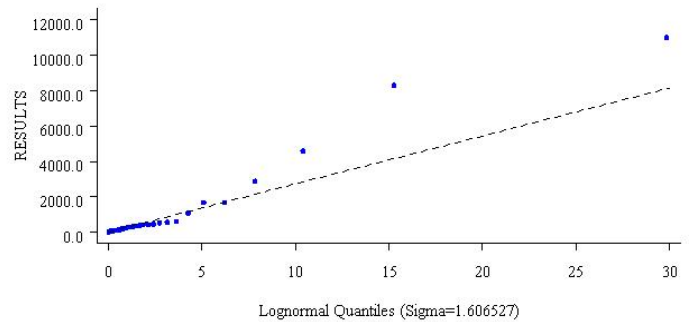
ENTACT, 1999

ENTACT, 1999

EPA / IDEM, 2002

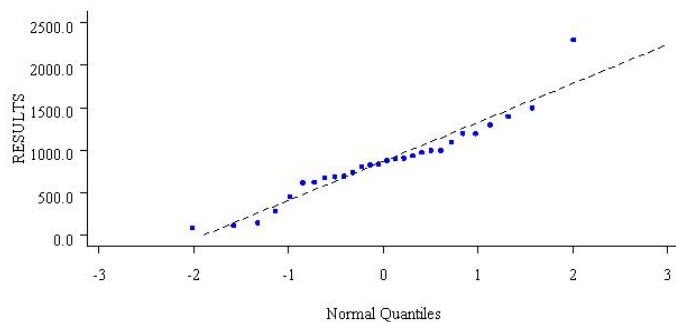### EPA / IDEM, 2002

Normal Q-Q Plots

Log-Normal Q-Q Plots

RCRA, 2003
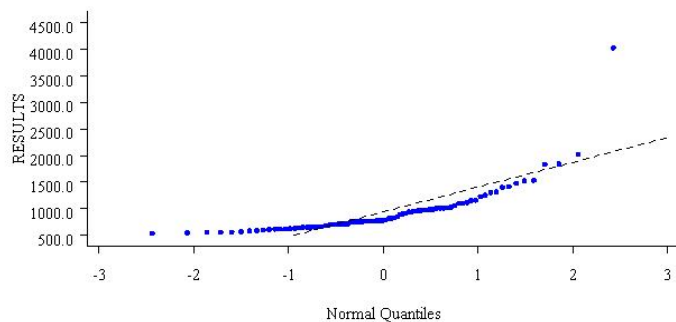
RCRA, 2003

EPA, 2006

EPA, 2006

250 m

500 m

750 m

1000 m

F

0	1,000	2,000

Feet

0	310	620

Meters

D:\GIS\Projects\USS_LEAD\MAPS\Total_Samples.pdf

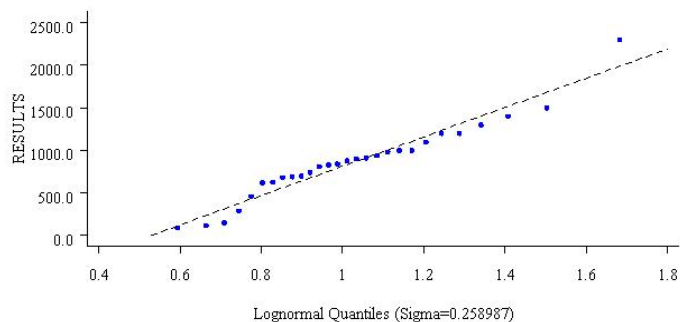| Surface Soil Lead (Pb) Concentration (ppm)	EPA 1985	ENTACT 1999	EPA / IDEM 2002	RCRA 2003	EPA 2006
0 - 399 ppm	⊕	X	K	!	ʏ
400 - 799 ppm	⊕	X	K	!	ʏ
800 - 1199 ppm	O	X	K	!	ʏ
>= 1200 ppm	⊕	X	K	!	ʏ
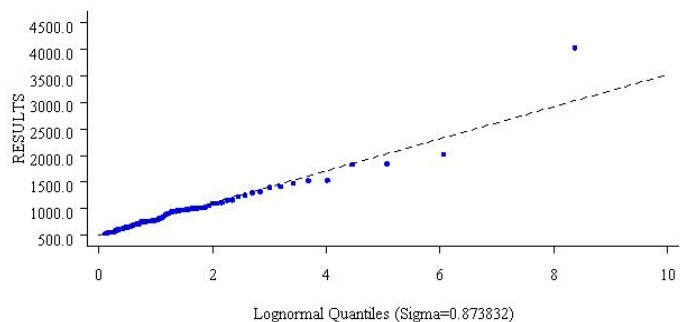
USS Lead Soil Contamination Site

East Chicago, IN

Surface Soil Samples Analyzed For Lead (Pb) Concentrations

⬚ USS Lead Site Boundary

FIELDS

UNITED STATES ENVIRONMENTAL PROTECTION AGENCY